k-means Approach to the Karhunen-Loéve Transform

Krzysztof Misztal, Przemysław Spurek and Jacek Tabor

***Abstract*—We present a simultaneous generalization of the well-known Karhunen-Loéve (PCA) and k-means algorithms. The basic idea lies in approximating the data with k affine subspaces of a given dimension n. In the case $n = 0$ we obtain the classical k-means, while for $k = 1$ we obtain PCA algorithm.**

We show that for some data exploration problems this method gives better result then either of the classical approaches.

***Index Terms*—Karhunen-Loéve Transform, PCA, k-Means, optimization, compression, data compression, image compression.**

I. INTRODUCTION

OUR general problem concerns splitting of a given dataset W into clusters with respect to their intrinsic dimensionality. The motivation to create such an algorithm is a desire to extract parts of data which can be easily described by a smaller number of parameters. More precisely, we want to find affine spaces $S_1, \ldots, S_k$ such that every element of W belongs (with certain maximal error) to one of the spaces $S_1, \ldots, S_k$.

To explain it graphically, let us consider following example. Figure 1(a) represents three lines in the plane, while Figure 1(b) a circle and an orthogonal line in the space. Our goal is to construct an algorithm that will split them into three lines and into a line and a circle.
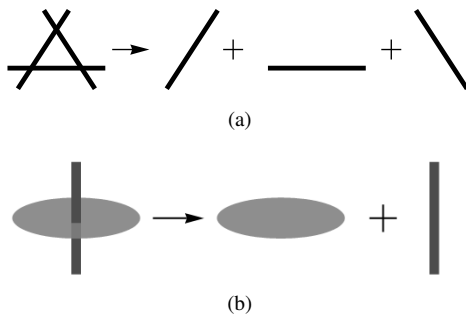

(a)

(b)

Fig. 1. Our goal is to create algorithm which will interpret Fig. 1(a) as three groups of one-dimensional points and Fig. 1(b) as two groups of one- and two-dimensional points.

We have constructed a simultaneous generalization of the k-means method [1] and the Karhunen-Loéve transform (called also PCA – Principle Component Analysis) [2] – we call it (ω, k)-means. Instead of finding k centers which best represent

K. Misztal, P. Spurek and J. Tabor are with Institute of Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. e-mails: krzysztof.misztal@ii.uj.edu.pl; przemyslaw.spurek@ii.uj.edu.pl; jacek.tabor@ii.uj.edu.pl.
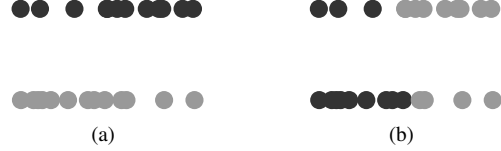
Manuscript received August 19, 2011.

(a)   (b)

Fig. 2. Example of clustering: Fig. 2(a) for $k = 2$ clusters which are 1–dimensional; Fig. 2(b) for classical k-means with $k = 2$ clusters.

the data as in the classical k-means, we find k nearest subspaces of a given fixed dimension n, where ω denotes the weight which takes part in measuring the distance. In analogy to the case of k-means, we obtain a version of the Voronoi diagram (for details see next section). In the simplest form our algorithm needs the number of clusters k and the dimension n (for $n = 0$ we obtain the k-means while for $k = 1$ we obtain the PCA).

To present our method, consider the points grouped along two parallel lines – Figure 2 presents the result of our program on the clustering of this set.

The approach can be clearly used in most standard applications of either the k-means or the Karhunen-Loéve transform. In particular, (since it is a generalization of the Karhunen-Loéve transform [3]) one of the possible natural applications of our method lies in the image compression. Figure 3 presents error[1] in image reconstruction of a classical Lena photo (508×508 pixels) as a function of k. Observe that just by modifying the number of clusters from 1 to 3, which makes the minimal increase in the necessary memory, we decrease twice the level of error in the compression.

Except for image compression our method can by applied in various situations where the classical k-means or PCA where used, for example in:

- data mining – we can detect important coordinates and subsets with similar properties;
- clustering – our modification of k-means can detect different, high dimensional relation in data;
- image compression and image segmentation;
- pattern recognition – thanks to detection of relation in data we can use it to assign data to defined before classes.

The basic idea of (ω, k)-means algorithm can be described as follows:

choose
    initial clusters distribution
repeat

---

[1] By error in image comparison we understand pixel by pixel image compare using standard Euclidean norm.
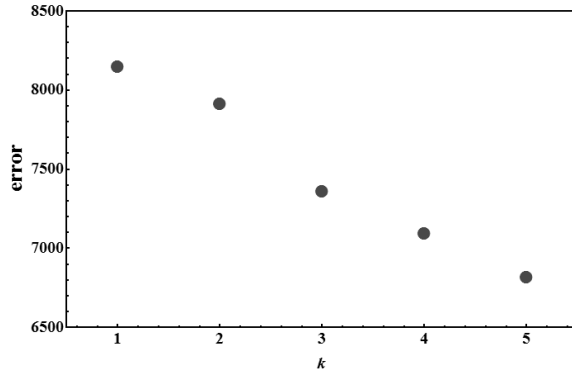
Fig. 3. Error in image decompression as a function of number of clusters k for $n = 5$.
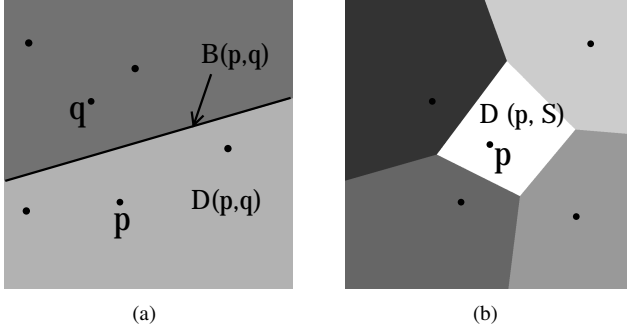

Fig. 4. Graphical presentation of $B(p,q)$, $D(p,q)$ and $D(p,S)$ in $\mathbb{R}^2$.

apply the Karhunen-Loéve method for each cluster
appoint new clusters
until decrease of "energy" is below given error

II. GENERALIZED VORONOI DIAGRAM

The Voronoi diagram is one of the most useful data structures in computational geometry, with applications in many areas of science [4]. For the convenience of the reader and to establish the notation we shortly describe the classical version of the Voronoi diagram (for more details see [5]). For $N \in \mathbb{N}$ consider $\mathbb{R}^N$ with the standard Euclidean distance and let S be a finite set of $\mathbb{R}^N$. For $p, q \in S$ such, that $p \neq q$, let

$$B(p,q) = \{z \in \mathbb{R}^N : \|p - z\| = \|q - z\|\}, \qquad (1)$$

$$D(p,q) = \{z \in \mathbb{R}^N : \|p - z\| < \|q - z\|\}. \qquad (2)$$

Hyperplane $B(p,q)$ divides $\mathbb{R}^N$ into two sets, one containing points which are closer to point p then q ($D(p,q)$), and the second one containing points which are closer to point q then p ($D(q,p)$) – see Figure 4(a).

Definition 2.1 ([5]): The set

$$D(p,S) := \bigcap_{q \in S : q \neq p} D(p,q)$$

of all points that are closer to p than to any other element of S is called the (open) Voronoi region of p with respect to S.

For $N = 2$ set $D(p,S)$ is the interior of a convex, possibly unbounded polygon (Figure 4(b)).

The points on the contour of $D(p,S)$ are those that have more than one nearest neighbor in S, one of which is p.

Definition 2.2 ([5]): The union

$$V(S) := \bigcup \partial D(p, S)$$

of all region boundaries is called the Voronoi diagram of S.

The common boundary of two Voronoi regions is a Voronoi edge. Two edges meet at a Voronoi vertex such a point has three or more nearest neighbors in the set S.

Now we proceed to the description of our modification of the Voronoi diagram. We divide the space $\mathbb{R}^N$ with respect to affine subspaces of $\mathbb{R}^N$.

Definition 2.3: For $n \leq N$ let

$$\mathrm{E}_n(\mathbb{R}^N) := \{(\mathrm{v}_0, \dots, \mathrm{v}_n) \in (\mathbb{R}^N)^{n+1} \text{ such that }$$
$$\mathrm{v}_i, \mathrm{v}_j \text{ are orthonormal for } i, j > 0, i \neq j\}.$$

Thus v_0 denotes a center of affine space we consider, while $\mathrm{v}_1, \dots, \mathrm{v}_n$ is the orthonormal base of its "vector part". From the geometrical point of view the element $\mathrm{v} = (\mathrm{v}_0, \mathrm{v}_1, \dots, \mathrm{v}_n) \in \mathrm{E}_n(\mathbb{R}^N)$ represents the affine space

$$\mathrm{v}_0 + \mathrm{lin}(\mathrm{v}_1, \dots, \mathrm{v}_n) = \mathrm{aff}(\mathrm{v}_0, \mathrm{v}_1, \dots, \mathrm{v}_n).$$

We modify equations (1) and (2), by using distance between a point and affine subspace generated by linear independence vector instead of distance between points.

Definition 2.4: Let $n < N$ and let $\mathrm{v} \in \mathrm{E}_n(\mathbb{R}^N)$, $\omega = (\omega_0, \dots, \omega_n) \in [0,1]^{n+1}$ such that $\sum\limits_{j=0}^{n} \omega_j = 1$ be given. For $x \in \mathbb{R}^N$ let

$$\mathrm{DIST}_\omega(x; \mathrm{v}) := \left(\sum_{j=0}^{n} \omega_j \ \mathrm{dist}(x; \mathrm{aff}(\mathrm{v}_0, \dots, \mathrm{v}_j))^2 \right)^{1/2}, \qquad (3)$$

where $\mathrm{dist}(x; V)$ denotes the distance of the point x from the space V.

In formula (3), $\omega = (\omega_0, \dots, \omega_n)$ is interpreted as vector of weights, where ω_k denotes the weight of the affine subspace of dimension k. It is easy to notice, that DIST has following properties:

- for $\mathrm{v} \in \mathrm{E}_n(\mathbb{R}^N)$ and $\omega = (0, \dots, 0, 1) \in [0,1]^{n+1}$ we obtain that $\mathrm{DIST}_\omega(x; \mathrm{v})$ is a distance between the point x and affine space $\mathrm{aff}(\mathrm{v})$;
- if $\mathrm{v}_0 = 0$ and $\omega = (0, \dots, 0, 1)$ then DIST_ω is a distance between point and linear space generated by $(\mathrm{v}_1, \dots, \mathrm{v}_n)$;
- if $\omega = (1, 0, \dots, 0)$ then DIST_ω is the classical distance between x and v_0:

$$\mathrm{DIST}_\omega(x; \mathrm{v}) = \|x - \mathrm{v}_0\|.$$

Remark 2.5: Formula (3) can be computed as follows

$$(\mathrm{DIST}_\omega(x; \mathrm{v}))^2 = \sum_{j=0}^{n} \omega_j \left(\|x - \mathrm{v}_0\|^2 - \sum_{i=1}^{j} \langle x - \mathrm{v}_0; \mathrm{v}_i \rangle^2 \right)$$
$$= \sum_{j=0}^{n} \omega_j \|x - \mathrm{v}_0\|^2 - \sum_{j=0}^{n} \omega_j \sum_{i=1}^{j} \langle x - \mathrm{v}_0; \mathrm{v}_i \rangle^2.$$
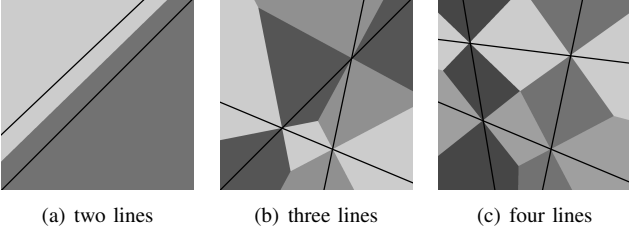
(a) two lines   (b) three lines   (c) four lines

Fig. 5. Generalized Voronoi diagram for $\omega = (0,1)$ and two, three and four lines on plane.

To optimize calculations we define

$$\bar{v}_1 = \langle x - v_0; v_1 \rangle^2,$$
$$\bar{v}_j = \bar{v}_{j-1} + \langle x - v_0; v_j \rangle^2,$$

and since $\sum \omega_j = 1$ thus we simplify our computation to

$$(\text{DIST}_\omega(x;v))^2 = \|x - v_0\|^2 - \sum_{j=0}^{n} \omega_j \bar{v}_j.$$

Now we are ready to define our generalization of the Voronoi diagram. Let S be a finite subset of $E_n(\mathbb{R}^N)$ and $\omega \in [0,1]^{n+1}$, $\sum \omega_j = 1$, where $n \leq N$. For $p,q \in S$ such, that $p \neq q$, let

$$B_\omega(p,q) := \{z \in \mathbb{R}^N : \text{DIST}_\omega(z;p) = \text{DIST}_\omega(z;q)\},$$

$$D_\omega(p,q) := \{z \in \mathbb{R}^N : \text{DIST}_\omega(z;p) < \text{DIST}_\omega(z;q)\}.$$

The set $B_\omega(p,q)$ divides the space $\mathbb{R}^N$ into two sets, first containing points which are closer to p then to q ($D_\omega(p,q)$) and second contain points which are closer to q then p ($D_\omega(q,p)$).

Definition 2.6: Let $n \in \mathbb{N}$, $n < N$ be fixed. Let S be finite subset of $E_n(\mathbb{R}^N)$ and $\omega \in [0,1]^{n+1}$, $\sum_{j=0}^{n} \omega_j = 1$ be given. For $p \in S$ the set

$$D_\omega(p,S) := \bigcap_{q \in S: q \neq p} D_\omega(p,q)$$

of all points that are closer to p than to any other element of S is called the (open) generalized Voronoi region of p with respect to S.

Applying this definition we obtain a new type of Voronoi diagram. As we can see in Figure 5, if $\omega = (0,1)$ we divide the plane into (not necessarily convex) polygons (similar situation to the classical Voronoi diagram). Figure 6 presents a generalized diagram on the plane for different weights changing from $\omega = (1,0)$ to $\omega = (0,1)$. In general we obtain that the boundary sets usually are not polygons but zeros of quadratic polynomials. The same happens in $\mathbb{R}^3$ even for $\omega = (0,1)$ see the Figure 7, where we show points with equal distance from two lines.

III. GENERALIZATION OF THE k-MEANS METHOD

Clustering is a classical problem of the division of the set $S \subset \mathbb{R}^N$ into separate clusters, or in other words, into sets showing given type of behavior.
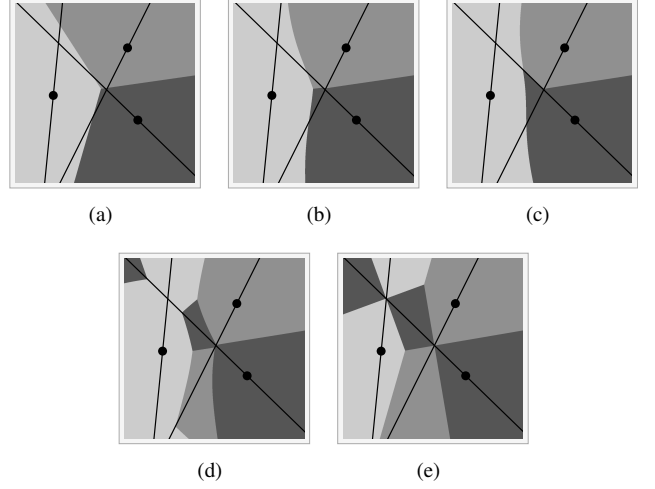

(a)   (b)   (c)

(d)   (e)

Fig. 6. Generalized Voronoi diagram for clustering of 3 clusters for different weight vectors Fig. 6(a), $\omega = (1,0)$; Fig. 6(b), $\omega = (\frac{3}{4}, \frac{1}{4})$; Fig. 6(c), $\omega = (\frac{1}{2}, \frac{1}{2})$; Fig. 6(d), $\omega = (\frac{1}{4}, \frac{3}{4})$; Fig. 6(e), $\omega = (0,1)$.
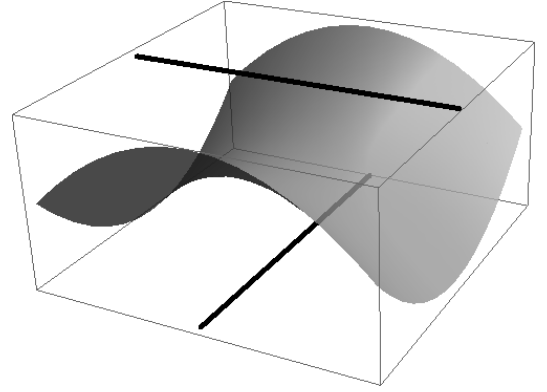

Fig. 7. Generalized Voronoi diagram for $\omega = (0,1)$ and two lines.

A. k-means

One of the most popular and basic method of clustering is the k-means algorithm. By this approach we want to divide S into k clusters $S_1, \ldots, S_k$ with minimal energy. For convenience of the reader and to establish the notation we shortly present the k-means algorithm.

For a cluster S and $r \in \mathbb{R}^N$ we define

$$E(S,r) := \sum_{s \in S} \|s - r\|^2.$$

The function $E(S,r)$ is often interpreted as an energy. We say that the point $\bar{r}$ best "describe" the set S if the energy is minimal, more precisely, if

$$E(S,\bar{r}) = \inf_{r \in \mathbb{R}^N} \{E(S,r)\}.$$

It is easy to show that barycenter (mean) of S minimizes the function $E(S,\cdot)$ (for more information see [6], [7]). The above consideration can be precisely formulated as follows:

Theorem 3.1 (k-means): Let S be a finite subset of $\mathbb{R}^N$. We have

$$E(S,\mu(S)) = \inf_{r \in \mathbb{R}^N} \{E(S,r)\}$$

where $\mu(S) := \frac{1}{\operatorname{card} S} \sum_{s \in S} s$ denotes the barycenter of S. Thus in the k-means the goal is to find such clustering $S = S_1 \cup \ldots \cup S_k$ that the function

$$E(S_1, \ldots, S_k) = \sum_{j=1}^{k} E(S_j, \mu(S_j))$$

is minimal. k-means algorithm for the set S after S. Lloyd [8]–[10] proceeds as follows:

stop condition
    choose $\varepsilon > 0$
initial conditions
    choose randomly points $\{\bar{s}_1, \ldots, \bar{s}_k\} \subset S$
    obtain first clustering $(S_1, \ldots, S_k)$ by matching each of the point $s \in S$ to the cluster S_j specified by $\bar{s}_j$ such that $\|s - \bar{s}_j\|^2$ is minimal
repeat
    let $E = E(S_1, \ldots, S_k)$
    compute new points $\bar{s}_1, \ldots, \bar{s}_k$ which best "describe" the clusters ($\bar{s}_j = \mu(S_j)$ for $j = 1, \ldots, k$)
    obtain new clustering $(S_1, \ldots, S_k)$ by adding each of the points $s \in S$ to the cluster such that $\|s - \bar{s}_j\|^2$ is minimal
until $E - E(S_1, \ldots, S_k) < \varepsilon$

Lloyd's method guarantees a decrease in each iteration but does not guarantee that the result will be optimal.

B. (ω, k)-means

In this chapter we consider generalization of k-means similar to that from the previous section concerning the Voronoi diagram. Instead of looking for the points which best "describe" clusters we seek n dimensional subspaces of $\mathbb{R}^N$.

Let $S \subset \mathbb{R}^N$ and $\omega \in [0,1]^{n+1}$, $\sum \omega_j = 1$ be fixed. For $v \in E_n(\mathbb{R}^N)$ let

$$E_\omega(S, v) := \sum_{s \in S} \operatorname{DIST}^2_\omega(s, v).$$

We interpret the function $E_\omega(S, v)$ as an energy of the set S respectively to the subspace generated by v. If the energy is zero, the set S is subset of affine space generated by v. We say that $\bar{v}$ best "describes" the set S if the energy is minimal, more precisely if

$$E_\omega(S, \bar{v}) = \inf_{v \in E_n(\mathbb{R}^N)} \{E_\omega(S, v)\}.$$

To obtain an optimal base we use a classical Karhunen-Loéve transform (called also Principal Component Analysis, shortly PCA), see [2]. The basic idea behind the PCA is to find the coordinate system in which the first few coordinates give us a "largest" possible information about our data.

Theorem 3.2 (PCA): Let $S = \{s_1, \ldots, s_m\}$ be a finite subset of $\mathbb{R}^N$. Let

$$\mathcal{M}(S) := (v_0, \ldots, v_N) \in E_N(\mathbb{R}^N)$$

be such that
- $v_0 = \mu(S)$;

- $v_1, \ldots, v_N$ are pairwise orthogonal eigenvectors of $[s_1 - v_0, \ldots, s_m - v_0] \cdot [s_1 - v_0, \ldots, s_m - v_0]^T$ arranged in descending order (according to the eigenvalues)[2].

For every $n < N$ and $\omega \in [0,1]^{n+1}$ we have

$$E_\omega(S, \mathcal{M}_k(S)) = \inf_{v \in E_n(\mathbb{R}^N)} \{E_\omega(S, v)\},$$

where $\mathcal{M}_k(S) := (v_0, \ldots, v_k)$.
Thus given $\omega \in [0,1]^{n+1}$, $\sum \omega_j = 1$, in (w, k)-means our goal is to find such clustering $S = S_1 \cup \ldots \cup S_k$ that the function

$$E_\omega(S_1, \ldots, S_k) := \sum_{j=1}^{k} E_\omega(S_j, \mathcal{M}_n(S)) \qquad (4)$$

is minimal. Consequently (ω, k)-means algorithm can be described as follows:

stop condition
    choose $\varepsilon > 0$
initial conditions
    choose randomly points $\{\bar{s}_1, \ldots, \bar{s}_k\} \subset S$
    obtain first clustering $(S_1, \ldots, S_k)$ by matching each of the points $s \in S$ to the cluster such that $\|s - \bar{s}_j\|^2$ is minimal
repeat
    let $E = E_\omega(S_1, \ldots, S_k)$
    compute vectors $v_1, \ldots, v_k$, which best "describe" the clusters, by the PCA method ($v_j = \mathcal{M}_n(S_j)$)
    obtain new clustering $(S_1, \ldots, S_k)$ by adding each of the point $s \in S$ to the cluster such that $\operatorname{DIST}_\omega(s, v_j)$ is minimal
until $E - E_\omega(S_1, \ldots, S_k) < \varepsilon$

As is the case in the classical k-means, our algorithm guarantees a decrease in each iteration but does not guarantee that the result will be optimal (cf. Example 3.3).

Example 3.3: As already mentioned in Section 2, the k-means do not find a global minimum and strongly depends on initial selection of clusters. In our case, this effect can be even more visible. Consider the case of circle C in $\mathbb{R}^2$ with 4 clusters and $\omega = (0, 1)$. The picture, see Figure 8(a), shows clustering obtained by use (ω, k)-means algorithm. Of course it is a local minimum of E_ω^C, however as we see at Figure 8(b) it is far from being the global minimum.

Initial cluster selection in our algorithm is the same as in k-mean algorithm, but it is possible to consider others ways:
- k-means++ algorithm [11];
- starting from a given division (not from random distribution);
- repeating the initial choice of clusters many times.

Each of above approaches usually solves the problem described in Example 3.3.

Remark 3.4: Let $S \subset \mathbb{R}^N$ and $v \in E_n(\mathbb{R}^N)$. It is easy to notice that the above method has following properties:
- for $\omega = (1, 0, \ldots, 0)$ we obtain the classical k-means,

---

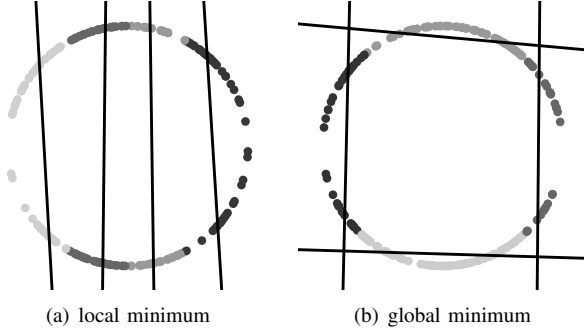[2]$[s_1 - v_0, \ldots, s_m - v_0]$ is a matrix with columns $s_j - v_0$, for $j = 1, \ldots, m$.

(a) local minimum   (b) global minimum

Fig. 8. Circle clustering in $\mathbb{R}^2$ for 4 clusters with $\omega = (0, 1)$. (ω, k)-means method strongly dependents on initial conditions.
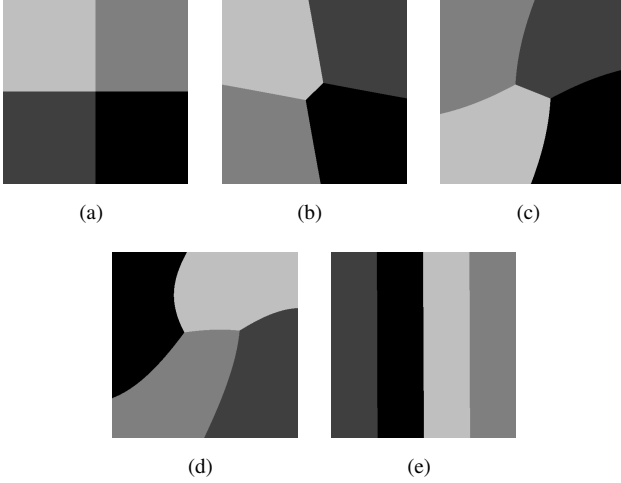

(a)   (b)

Fig. 10. Clustering with: Fig. 10(a) – k-means; Fig. 10(b) – (ω, k)-means.

(a)   (b)   (c)

(d)   (e)

Fig. 9. (ω, k)-means method for clustering into 4 clusters of set $\{\frac{0}{1000}, \frac{1}{1000}, \ldots, \frac{1000}{1000}\} \times \{\frac{0}{1000}, \frac{1}{1000}, \ldots, \frac{1000}{1000}\}$ for different weight vectors: Fig. 6(e), $\omega = (1, 0)$; Fig. 6(d), $\omega = (\frac{3}{4}, \frac{1}{4})$; Fig. 6(c), $\omega = (\frac{1}{2}, \frac{1}{2})$; Fig. 6(b), $\omega = (\frac{1}{4}, \frac{3}{4})$; Fig. 6(a), $\omega = (0, 1)$.

- for $n = 1$ we get Karhunen-Loéve transform.

As an algorithm's outcome we get:
- division of the data into clusters $\{S_1, \ldots, S_k\}$;
- for each cluster an affine space of dimension n obtained by the Karhunen-Loéve method which best represents the given cluster.

Example 3.5: If we apply our algorithm for regular plane subset (ex. square) we obtain generalized Voronoi diagram (cf. Fig. 6) – Figure 9 present clustering for different weight vector changing from $\omega = (1, 0)$ to $\omega = (0, 1)$.

IV. APPLICATIONS

A. Clustering

Clustering, by (ω, k)-means algorithm, gives a better description of the internal geometry of a set, in particular it found a reasonable splitting into connected components of consider the points grouped along two parallel sections (see Figure 2). Similar effect we can see in next example, when we consider the points grouped along circle and interval, see Figure 10.

Concluding, in many cases the (ω, k)-means method can be very useful in seeking n-dimensional (connected) components of given data sets.
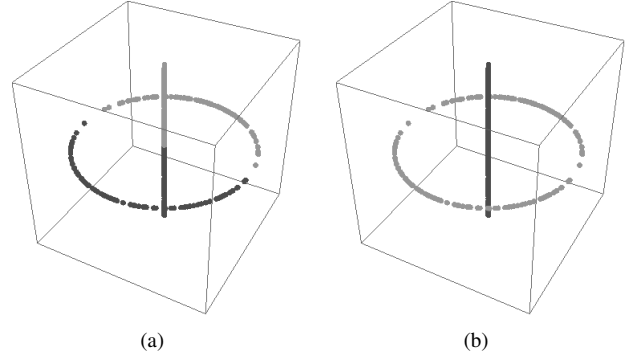
B. Analysis of Functions

In this subsection we consider real data from acoustics. Acoustical engineers [12] study reverberation which is observed when a sound is produced in an enclosed space causing a large number of echoes to build up and then slowly decay as the sound is absorbed by the walls and the air. Reverberation time is crucial for describing the acoustic quality of a room or space. It is the most important parameter for describing sound levels, speech intelligibility and the perception of music and is used to correct or normalize building acoustics and sound power measurements.

We analyze the decay curve (see Figure 11(a)) which presents measurement of sound level in time and describe way in which sound impulse vanishing into background noise. Based on this we want to recover reverberation time. In particular we know that we have two linear component: first connected with sound absorption by the space and second – background noise. To use statistical analysis, we have to extract both of them. Our algorithm detects the n-dimensional subspaces, so for the parameters $\omega = (0, 1)$ we obtain linear approximation of function, see Figure 11(b).

Results obtained using our algorithm are comparable with those obtained by classical methods and give more opportunities for further research.

C. Image compression

Our algorithm can be used to compress images. First, we interpret photo as a matrix. We do this by dividing it into 8 by 8 pixels, where each pixel is described (in RGB) by using 3 parameters. Each of the pieces is presented as a vector from $\mathbb{R}^{192}$. By this operation we obtain dataset from $\mathbb{R}^{192}$.

Taking into consideration the classical Lena picture (508×508 pixels), let us present its compressed version with the use of k-means method (Figure 12(a), $k = 5$, $n = 0$), Karhunen-Loéve Transform (Figure 12(b), $k = 1$, $n = 1$) and (ω, k)-means algorithm (Figures 12(c) and 12(d)). As we can see the algorithm allows to reconstruct with great accuracy compressed images while reducing the amount of needed information to save (in our example we remember ex. only 5 coordinates in 192-dimensional space).

Table I presents error in image reconstruction for Lena picture. We run (ω, k)-means algorithm 16 times and each run improve clustering quality 50 times.
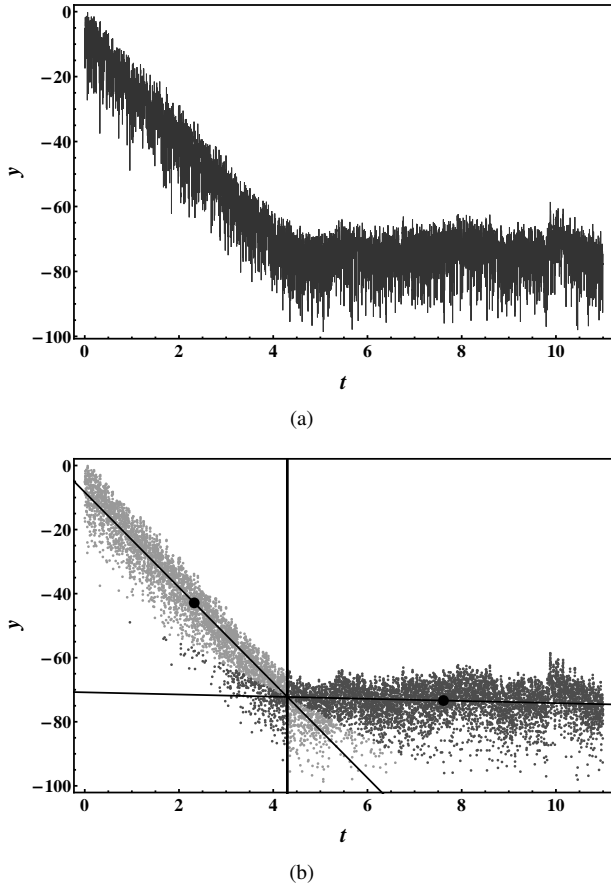
(a)

(b)

Fig. 11. Linear component of the data structure. Fig. 11(a) – decay curve (original data). Fig. 11(b) – outcome from (ω, k)-means algorithm for $k = 2$, $\omega = (0, 1)$ we extract two linear components in data (black dots match clusters centers with the corresponding lines describing those clusters, vertical line separate sound and background noise – after 4.3 s).

TABLE I
ERROR IN IMAGE DECOMPRESSION FOR CERTAIN k AND n.

k	n					
	0	1	2	3	4	5
1	40328.6	19499.3	16358.1	12452.1	10160.5	8149.7
2	27502.8	17193.3	13031.9	10382.8	9082.7	7913.2
3	23261.2	15437.2	11631.8	9612.0	8350.1	7358.3
4	20990.4	14454.1	11004.1	9192.7	7922.4	7095.9
5	20150.1	13740.0	10602.4	8867.2	7745.9	6814.5

V. IMPLEMENTATION

Sample implementation of (ω, k)-means algorithm prepared in Java programming language is available at http://www.ii.uj.edu.pl/~misztalk.

REFERENCES

[1] A. Jain and R. Dubes, *Algorithms for clustering data*.   Prentice Hall, 1988.
[2] I. Jolliffe, "*Principal component analysis*," *Encyclopedia of Statistics in Behavioral Science*, 2002.
[3] D. Salomon, *Data compression: the complete referenc*. Springer-Verlag New York Inc, 2007.
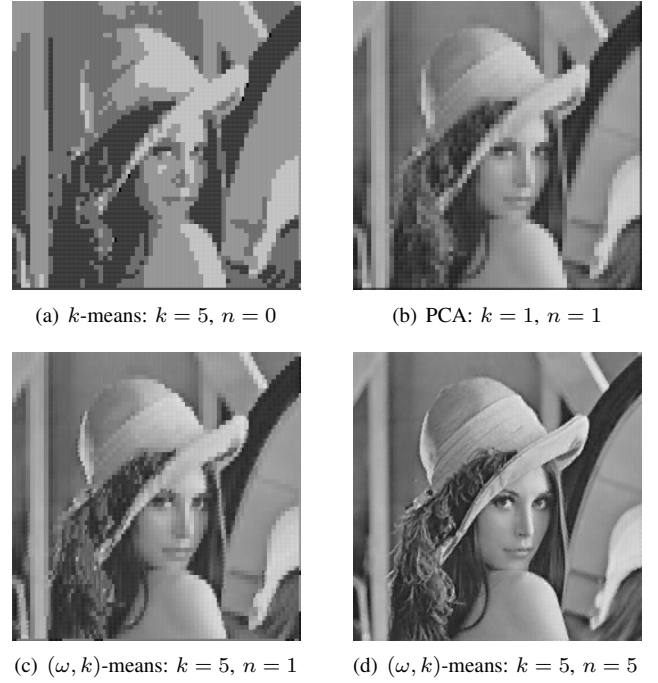[4] A. Okabe, *Spatial tessellations: concepts and applications of Voronoi diagrams*.   John Wiley & Sons Inc, 2000.

(a) k-means: $k = 5$, $n = 0$

(b) PCA: $k = 1$, $n = 1$

(c) (ω, k)-means: $k = 5$, $n = 1$

(d) (ω, k)-means: $k = 5$, $n = 5$

Fig. 12.   Compressed version of Lena picture. Subimage compare: Fig. 12(a) – classical k-means; Fig. 12(b) – Karhunen-Loéve Transform; Fig. 12(c) and Fig. 12(d) – (ω, k)-means algorithm.

[5] R. Klein, "*Concrete and Abstract Voronoi Diagrams, volume 400 of Lecture Notes Comput. Sci*," *Springer-Verlag*, vol. 26, p. 27, 1989.
[6] C. Ding and X. He, "*K-means clustering via principal component analysis*," in *Proceedings of the twenty-first international conference on Machine learning*.   ACM, 2004, p. 29.
[7] W. Fisher, "*On grouping for maximum homogeneity*," *Journal of the American Statistical Association*, vol. 53, no. 284, pp. 789–798, 1958.
[8] E. Forgy, "*Cluster analysis of multivariate data: Efficiency vs. interpretability of classification*," *Biometrics*, vol. 21, no. 3, pp. 768–769, 1965.
[9] S. Lloyd, "*Least squares quantization in PC*," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, 2002.
[10] J. MacQueen *et al.*, "*Some methods for classification and analysis of multivariate observation*," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, no. 281-297. California, USA, 1967, p. 14.
[11] D. Arthur and S. Vassilvitskii, "*k-means++: The advantages of careful seeding*," in *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*.   Society for Industrial and Applied Mathematics, 2007, pp. 1027–1035.
[12] T. Kamisiński, J. Rubacha, and A. Pilch, "*The Study of Sound Scattering Structures for the Purposes of Room Acoustic Enhancement*," *Acta Physica Polonica A, Polska Akademia Nauk. Instytut Fizyki, Warszawa*, vol. 118, no. 1, pp. 83–86, 2010.